Updated September 5, 2006

WOMBAT and WOMBAT-PK

Marius Olah, Maria Mracec, Ramona Rad, Liliana Ostopovici, Alina Bora, Nicoleta Hadaruga, Dan Hadaruga, Ramona Moldovan, Adriana Fulias, and Tudor I. Oprea

Sunset Molecular Discovery LLC
1704 B Llano St., S-te 140
Santa Fe, NM 87505-5415

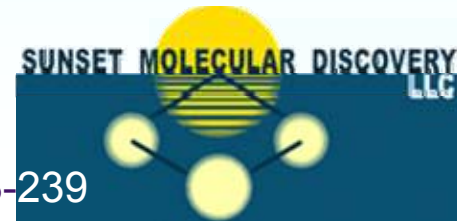
Contact Info:
tudor@sunsetmolecular.com
Tel: +1 505 550 7744

EuroQSAR 2006 Demo CD
Civitavecchia, Italy, September 10, 2006
© 2006 Tudor I. Oprea, all rights reserved

SUNSET MOLECULAR DISCOVERY LLC

WOMBAT 2006.1

- **WOMBAT 2006.1** contains **154,236 entries** (136,091 unique SMILES), totaling **307,700** biological activities on over **1,320** unique targets.
- **WOMBAT 2006.1** contains **6,801** different series from **6,791 papers** published in medicinal chemistry journals between 1975 and 2005
- *Systematic coverage for:* ***J. Med. Chem.*** (77.6%) **1991-2004** [complete], 2005 [partial], ***Bioorg. Med. Chem. Lett.*** (15.4%) **2002-2003** [complete], 2004 [partial], **Bioorg. Med. Chem.** (5.6%) **2002-2003** [complete], ***Eur. J. Med. Chem.*** (1%) **2002-2003** [complete] 2004 [partial];
- **SwissProt** IDs for ~88% of the Entries
- DOI (digital object identifier) links & PubMed IDs for all references (direct access to PDF files for institutions with appropriate subscriptions)
- ClogP and XMR from Biobyte Corporation (Al Leo), AlogP and LogSw from ALOGPS (Igor Tetko), Ligand Efficiency, Rule-of-Five, and Molecular Complexity can be queried.
- Over 10,000 unique entries are added every six months



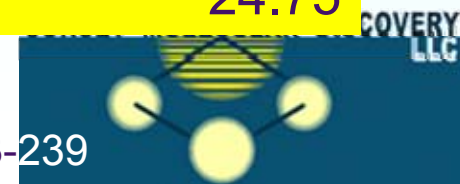
M. Olah et al. *Chemoinformatics in Drug Discovery,* Wiley-VCH **2005,** pp 223-239

Activity Profile for WOMBAT 2006.1

Target Class	Compounds [a]	Percentage
G-protein coupled receptors (GPCRs)	50,791	32.93
Integrins	3,127	2.03
Nuclear Hormone Receptors	4,335	2.81
Other Receptors	3,375	2.59
Ion Channels	13,500	8.75
Serine Proteases	7,596	7.80
(Oxido)Reductases	7,770	5.19
Kinases	9,705	5.95
Oxygenases	6,051	3.77
Aspartyl Proteases	4,904	3.22
Serine Proteases	7,596	7.80
Metallo-proteases	4,296	1.53
Transporters	5,462	3.54
Others	38,170	24.75

(a) number of structures active at least once/target, % of total entries

M. Olah et al. *Chemoinformatics in Drug Discovery,* Wiley-VCH **2005,** pp 223-239



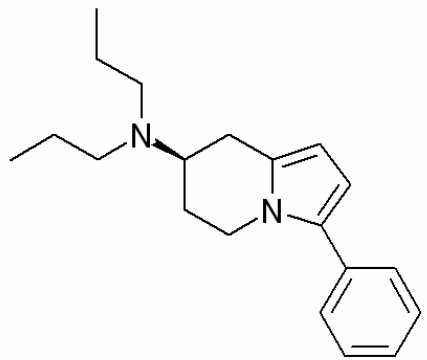
Bioactivity Summary | Target & Biological Information | Computed Chemical Properties

Sunset Molecular Discovery, LLC
www.sunsetmolecular.com

W O M B A T

SMDL-00150472	6641

structure



CCCN(CCC)[C@@H]1CCn2c(ccc2C1)c3ccccc3

SMDL_ID	molname
150472	(R)-28
generic_name	
reference	J. Med. Chem. 48(7)-2005 2493-2508
formula	mol_weight
$C_{20}H_{28}N_2$	296.4596
ALogP	ExpLogP
5.3100	
ALogS	ExpLogS
-3.9600	

structure_keywords
abs
-

act#	targ_type	target_name	act_type	act_val	act_min	act_max	inhib%	lig_eff
1	receptor	D1	Ki low	<4.6990				
3	receptor	D2	Ki low	4.9208	4.9208	4.9208		0.3176
9	receptor	D3	Ki high	5.6200	5.6200	5.6200		0.3627
10	receptor	D3	Ki low	5.1024	5.0590	5.1506		0.3293

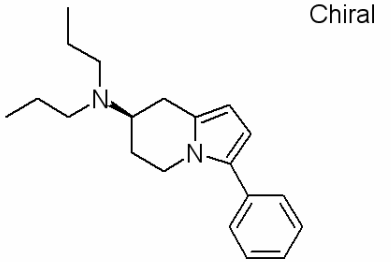
Olah et al., WOMBAT & WOMBAT-PK – for Wiley-VCH bookchapter 2006

Bioactivity Summary | **Target & Biological Information** | Computed Chemical Properties

Sunset Molecular Discovery, LLC
www.sunsetmolecular.com

W O M B A T

SMDL-00150472	6641



act#	target name(s)	activity value
3	D2; D2 dopamine receptor	4.9208
9	D3; D3 dopamine receptor	5.6200
10	D3; D3 dopamine receptor	5.1024

SwissProt

ID	P35462; DRD3_HUMAN
species	human

Protein

family	
type	

Biology

species	human
tiss_type	membrane
tiss_src	cloned
cell	CHO cell
i_cell_str	
stimulus	
effect	antagonist

Enzyme

family	
type	
EC	
substrate	
reactant	
cofactor	
radiosub	

Binding

endog_lig	dopamine
radio_lig	[3H]spiperone
nonspec	haloperidol

Other

actdet	

Receptor

family	GPCR
type	

GPCR class:

#1	Class A Rhodopsin like
#2	Amine
#3	Dopamine
#4	Dopamine Vertebrate type 3
#5	

NHR class:

#1	
#2	
#3	

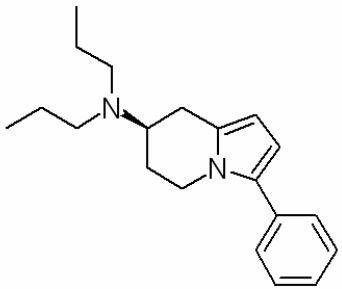
DISCLAIMER: Family and subfamily classifications for receptors was extracted from the following web-based databases: http://www.gpcr.org/7tm/ for GPCRs, and http://www.receptors.org/N for NHRs. This data is provided free-of-charge, as-is, for informative purposes only. Sunset Molecular Discovery LLC assumes no responsibility for any inaccuracies or errors that may be present in the target classification fields, related, but not limited to, SwissProt ID, Enzyme Codebook number or family classification.

Bioactivity Summary | Target & Biological Information | **Computed Chemical Properties**

Sunset Molecular Discovery, LLC
www.sunsetmolecular.com

W O M B A T

SMDL-00150472	6641



Number of ...

atoms	22	2	heteroatoms
bonds	24	18	non-polar atoms
rings	3	1	positive ioniz.
rotatable bonds	7	0	negative ioniz.
rigid bonds	15		

Lipinski RO5

HB-donors	0
HB-acceptors	0
cLogP	5.3460
mol weight	296.4596
total score	1

Surface area

polar	1.7469
non-polar	387.6991
%-polar	0.0045
%-non-polar	0.9955

ALogPS

ALogP	5.3100
ExpLogP	
ALogS	-3.9600
ExpLogS	

Others

XMR	1.1280
ABE	9.5000
SMCM	49.6810

Ligand efficiency

act#	activity	lig_eff
9	5.6200	0.3627
10	5.1024	0.3293

Legend

XMR: Molar refractivity excess
ABE: Andrews binding energy
SMCM: Synthesis and molecular complexity metric

Olah et al., WOMBAT & WOMBAT-PK – for Wiley-VCH bookchapter 2006

WOMBAT-PK 2006

- 900+ drugs indexed from multiple literature sources: Goodman & Gilman's 9th Edition, 1996; Avery's 4th Edition, 1997; Physician Desk Reference 2001; FDA Approved Drug Labels; and peer-reviewed literature
- **Physico-chemical properties** captured from Hansch, Leo & Hoekman Tables ($LogD_{7.4}$; LogP) and Avery's (pKa)
- ClogP and XMR from Biobyte Corporation (Al Leo), AlogP and LogSw from ALOGPS (Igor Tetko), Ligand Efficiency, Rule-of-Five, and Molecular Complexity can be queried.
- **MRTD** (minimum recommended therapeutic daily dose) was collected from the CDER@FDA website
- 753 drugs have **Drug Targets** (97% with **SwissProt** IDs); 419 drugs have **phase I metabolism enzymes** (all with **SwissProt** IDs)
- **Cardiac Toxicity:** 218 drugs have QT-prolongation data; 89 are assessed for *Torsade de Pointes* risk; 71 drugs have hERG binding data
- **VD_{ss}:** 552 drugs (309 i.v.) have volume of distribution data.
- **CL:** 514 drugs (233 i.v.) have total clearance data.

T. Oprea et al. *Molecular Interaction Fields,* Wiley-VCH **2006,** 249-272

SUNSET MOLECULAR DISCOVERY LLC

Measured Data in WOMBAT-PK 2006

Property	On File	G & G	Avery
%Oral Bioavailability	740	312	277
%Urinary excretion	339	326	N/A
%Plasma Protein Binding	776	311	434
Clearance, Cl (mL/min*kg)	514	320	422
Nonrenal clearance (fractional)	442	N/A	442
Volume of distribution, VDss (L/kg)	552	322	453
Half-life, T1/2 (hrs)	839	338	576
Terminal half-life, TT1/2 (hrs)	581	N/A	580
Effective concentration (mM/L)	119	118	N/A
MRTD (µmole/kg-bw/day)	575	N/A	N/A
MRTD_U (µmole/kg-bw/day; f_u corrected)	498	N/A	N/A
LogD7.4 (measured)	513	N/A	N/A
LogPoct (measured)	472	N/A	N/A
pKa1	350	N/A	274
pKa2	99	N/A	75
In Vitro Binding Data (from WOMBAT)	453	N/A	N/A

T. Oprea et al. *Molecular Interaction Fields,* Wiley-VCH **2006,** 249-272

SUNSET MOLECULAR DISCOVERY LLC

Clinical Properties (Ranked)

%Oral	Rank3Oral	Rank5Oral
0-5	**0**	**0**
5.1-19.99	0	1
20.0-79.99	1	2
80.0-95	2	3
>95.1	2	4

%Urine	RankUrine
0-1	0
1.01-5	1
5.01-20	2
20.01-50	3
50.01-80	4
>80	5

%PPB	RankPPB	Cl (mL/min*kg)	RankCl	VD (L/kg)	RankVD
0-5	0	0-(6/7)	0	0-0.1714	0
5.01-20	1	(6.01/7)-(12/7)	1	0.17141-0.5	1
20.01-80	2	(12.01/7)-5	2	0.501-1.0	2
80.01-95	3	5.01-10	3	1.001-3	3
95.01-99	4	**10.01-15.5**	**4**	**3.001-10**	**4**
>99.1	**5**	**>15.5**	**5**	**>10**	**5**

- Data were curated using a systematic round-off procedure: e.g., “<5” was attributed a higher value (=2.5), compared to “<1” (=0.5).
- Incremental rank values (selected from experience) are used when possible: e.g., experimental errors related to %Oral occur mostly for values between 20 and 80%;
- “6/7” and “12/7” represent the ½ and full value of creatinine clearance (120 mL/70 kg/min), respectively – this represents the renal component of the total (incl. hepatic) clerance; etc.

T. Oprea et al. *Molecular Interaction Fields,* Wiley-VCH **2006,** 249-272

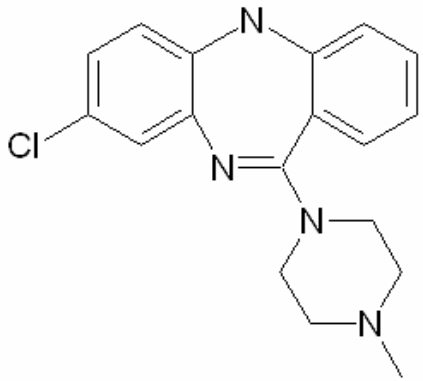
SUNSET MOLECULAR DISCOVERY LLC

Compound Description | Pharmacokinetic Data | Potential Side Effects | Computed Chemical Properties

Sunset Molecular Discovery, LLC
www.sunsetmolecular.com

W O M B A T - P K

ID	CAS#	molname
88	5786-21-0	clozapine



CN1CCN(CC1)C2=Nc3cc(Cl)ccc3Nc4ccccc24

FMLA	MW	stereo	mixture
$C_{18}H_{19}ClN_4$	326.8319	nc	

ExpLogP	EstLogP	ExpLogD7.4	EstLogD7.4
3.2300	3.8000	2.9900	

F	derivative	CAS#
0	//	5786-21-0

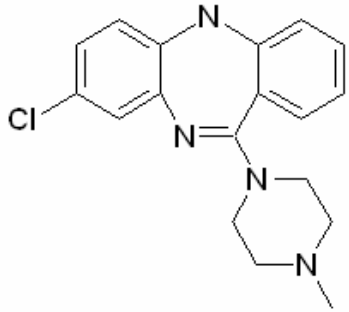
trade_name	
Clozaril (Novartis); Leponex (Novartis)	
drug_code	MP (*C)
HF-1854	183-184

ExpWSol	unit	comment

pKa1	pKa2	pKa3	pKa4
3.7000	7.6000		
pKa1_moiety	pKa2_moiety	pKa3_moiety	pKa4_moiety

Sunset Molecular Discovery, LLC
www.sunsetmolecular.com

W O M B A T - P K



ID	88
CAS#	5786-21-0
molname	clozapine
FMLA	$C_{18}H_{19}ClN_4$
MW	326.8319

drug_category	antipsychotic
comment	is a weak dopanime1 and dopamine2 receptor blocker
dosage forms	TABLET
biotransformation	Hepatic

PK data

oral_availability	%urin.excretion	%plasma bound	VDss [L/kg]	iv	CL [mL/min*kg]	iv
52.50	50.00	96.33	5.40		8.50	

eff.conc. [ng/mL]	T1/2 [hours]	T1/2 term [hours]	[Po]	nature
122.00	14.00	16.00		

notes (urinary exc.<1 in G&G; data above from FDA Label by Novartis)
Substrate for CYP 1A2, 2D6, 3A4

phase1 metab. enzyme	phase1 metab. swissprotID
CYP1A2	CP1A2_HUMAN
CYP2C8	CP2C8_HUMAN
CYP2C9	CP2C9_HUMAN
CYP2D6	CP2D6_HUMAN

Ranking & MRTD

Rank3Oral	Rank5Oral	RankUrine	RankPPB	RankCL	RankVD
1	2	4	4	3	4

mrtd [mg/kgbw/day]	mrtd [microM/day]	mrtd_U [microM/day]
15.00	45.90	1.6828

Drug Target

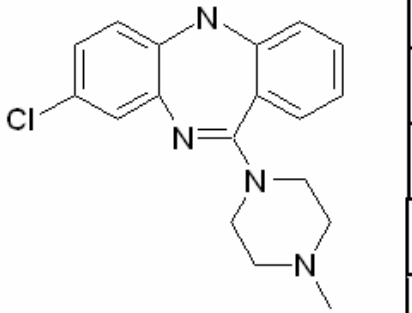
name	class	swissprotID	pdbID
D4 dopamine receptor	gpcr	DRD4_HUMAN (P21917)	1M2C
histamine H1 receptor	gpcr	HRH1_HUMAN (P35367)	
D2 dopamine receptor	gpcr	DRD2_HUMAN (P14416)	2CII

LEGEND:

- fields pk_vdss_admin & pk_clearance_admin: iv, po

Sunset Molecular Discovery, LLC
www.sunsetmolecular.com

W O M B A T - P K



ID	88
CAS#	5786-21-0
molname	clozapine
FMLA	$C_{18}H_{19}ClN_4$
MW	326.8319

LD50 / derivative: //

LD50 value	unit	type	species	comment
61.00	mg/kg	i.v.	mouse	
58.00	mg/kg	i.v.	rat	
199.00	mg/kg	oral	mouse	

WOMBAT bioactivities

reference	target type	target name	act type	act val	inhib%	lig eff
Bioorg. Med. Chem. Lett. 12(4)-2002 633-636	receptor	D1	Ki	6.3768		0.3937
Bioorg. Med. Chem. Lett. 12(4)-2002 633-636	receptor	D2L	Ki	7.3872		0.4561
Bioorg. Med. Chem. Lett. 12(4)-2002 633-636	receptor	D2S	Ki	7.5528		0.4663
Bioorg. Med. Chem. Lett. 12(4)-2002 633-636	receptor	D3	Ki	6.0177		0.3715
Bioorg. Med. Chem. Lett. 12(4)-2002 633-636	receptor	D4	Ki	7.7959		0.4813

BBB & TdP

BBB value	BBB+
TdP class	
TdP warnings	with-possible-risk; with-risk-for-congenital-lo

hERG bioactivities

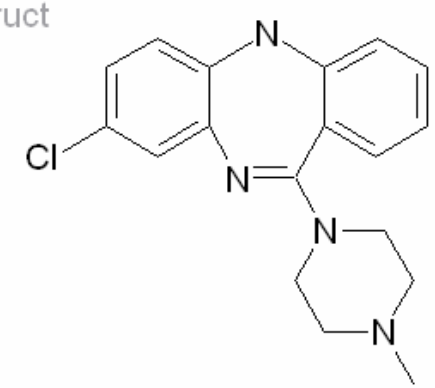
reference	act type	act val
Bioorg. Med. Chem. Lett. 13(16)-2003 2773-2775	IC50	6.4900
J. Chem. Inf. Model. 46(1)-2006 392-400	IC50	6.5900
J. Med. Chem. 45(18)-2002 3844-3853	IC50	6.7190

Sunset Molecular Discovery, LLC
www.sunsetmolecular.com

W O M B A T - P K

clozapine	5786-21-0

molstruct



Number of ...

atoms	23	4	heteroatoms
bonds	26	15	non-polar atoms
rings	4	2	positive ioniz.
rotatable bonds	2	0	negative ioniz.
rigid bonds	22		

Lipinski RO5

HB-donors	1
HB-acceptors	0
cLogP	3.71
mol weight	326.83
total score	0

Surface area

polar	25.6290
non-polar	324.6680
%-polar	0.0732
%-non-polar	0.9268

ALogPS

ALogP	3.8000
ExpLogP	3.2300
ALogS	-3.2300
ExpLogS	

Others

XMR	1.5950
ABE	14.8000
SMCM	50.5550

Legend

XMR: Molar refractivity excess
ABE: Andrews binding energy
SMCM: Synthesis and molecular complexity metric